Chapter 3: Elements of Probability

1. A box contains three marbles — one red, one green, and one blue. Consider an experiment that consists of taking one marble from the box, then replacing it in the box and drawing a second marble from the box. Describe the sample space.

Repeat for the case in which the second marble is drawn without first replacing the first marble.

2. An experiment consists of tossing a coin three times. What is the sample space of this experiment? Which event corresponds to the experiment resulting in more heads than tails?

3. Let $S = \{1, 2, 3, 4, 5, 6, 7\}$, $E = \{1, 3, 5, 7\}$, $F = \{7, 4, 6\}$, $G = \{1, 4\}$. Find

    (a) EF ;   **(c)** EG^c;   **(e)** $E^c(F \cup G)$;

    (b) $E \cup FG$;   **(d)** $EF^c \cup G$;   **(f)** $EG \cup FG$.

4. Two dice are thrown. Let E be the event that the sum of the dice is odd, let F be the event that the first die lands on 1, and let G be the event that the sum is 5. Describe the events EF, $E \cup F$, FG, EF^c, EFG.

5. A system is composed of four components, each of which is either working or failed. Consider an experiment that consists of observing the status of each component, and let the outcome of the experiment be given by the vector (x_1, x_2, x_3, x_4) where x_i is equal to 1 if component i is working and is equal to 0 if component i is failed.

    (a) How many outcomes are in the sample space of this experiment?

    (b) Suppose that the system will work if components 1 and 2 are both working, or if components 3 and 4 are both working. Specify all the outcomes in the event that the system works.

    (c) Let E be the event that components 1 and 3 are both failed. How many outcomes are contained in event E?

6. Let E, F, G be three events. Find expressions for the events that of E, F, G

    (a) only E occurs;

    (b) both E and G but not F occur;

    (c) at least one of the events occurs;

    (d) at least two of the events occur;

    (e) all three occur;

(f) none of the events occurs;

(g) at most one of them occurs;

(h) at most two of them occur;

(i) exactly two of them occur;

(j) at most three of them occur.

7. Find simple expressions for the events

(a) $E \cup E^c$;

(b) EE^c;

(c) $(E \cup F)(E \cup F^c)$;

(d) $(E \cup F)(E^c \cup F)(E \cup F^c)$;

(e) $(E \cup F)(F \cup G)$.

8. Use Venn diagrams (or any other method) to show that

(a) $EF \subset E$, $E \subset E \cup F$;

(b) if $E \subset F$ then $F^c \subset E^c$;

(c) the commutative laws are valid;

(d) the associative laws are valid;

(e) $F = FE \cup FE^c$;

(f) $E \cup F = E \cup E^c F$;

(g) DeMorgan's laws are valid.

9. For the following Venn diagram, describe in terms of E, F, and G the events denoted in the diagram by the Roman numerals I through VII.
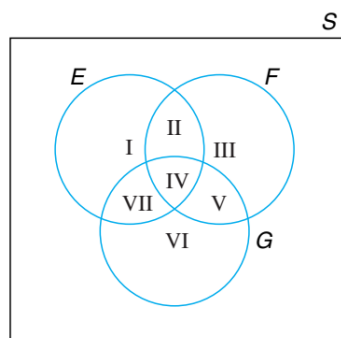
10. Show that if $E \subset F$ then $P(E) \leq P(F)$. (*Hint:* Write F as the union of two mutually exclusive events, one of them being E.)

11. Prove Boole's inequality, namely that

$$P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$$

12. If $P(E) = .9$ and $P(F) = .9$, show that $P(EF) \geq .8$. In general, prove Bonferroni's inequality, namely that

$$P(EF) \geq P(E) + P(F) - 1$$

13. Prove that

   (a) $P(EF^c) = P(E) - P(EF)$

   (b) $P(E^cF^c) = 1 - P(E) - P(F) + P(EF)$

14. Show that the probability that exactly one of the events E or F occurs is equal to $P(E) + P(F) - 2P(EF)$.

15. Calculate $9C3$, $9C6$, $7C2$, $7C5$, $10C7$.

16. Show that

$$nCr = nC(n - r)$$

Now present a combinatorial argument for the foregoing by explaining why a choice of r items from a set of size n is equivalent to a choice of $n-r$ items from that set.

17. Show that

$$nCr = (n-1)C(r-1) + (n-1)Cr$$

For a combinatorial argument, consider a set of n items and fix attention on one of these items. How many different sets of size r contain this item, and how many do not?

18. A group of 5 boys and 10 girls is lined up in random order — that is, each of the 15! permutations is assumed to be equally likely.

   (a) What is the probability that the person in the 4th position is a boy?

   (b) What about the person in the 12th position?

   (c) What is the probability that a particular boy is in the 3rd position?

19. Consider a set of 23 unrelated people. Because each pair of people shares the same birthday with probability 1/365, and there are $23C_2$ = 253 pairs, why isn't the probability that at least two people have the same birthday equal to 253/365?

20. Suppose that distinct integer values are written on each of 3 cards. These cards are then randomly given the designations A, B, and C. The values on cards A and B are then compared. If the smaller of these values is then compared with the value on card C, what is the probability that it is also smaller than the value on card C?

21. There is a 60 percent chance that the event A will occur. If A does not occur, then there is a 10 percent chance that B will occur.

  (a) What is the probability that at least one of the events A or B occurs?

  (b) If A is the event that the democratic candidate wins the presidential election in 2012 and B is the event that there is a 6.2 or higher earthquake in Los Angeles sometime in 2013, what would you take as the probability that both A and B occur? What assumption are you making?

22. The sample mean of the annual salaries of a group of 100 accountants who work at a large accounting firm is $130,000 with a sample standard deviation of $20,000. If a member of this group is randomly chosen, what can we say about

  (a) the probability that his or her salary is between $90,000 and $170,000?

  (b) the probability that his or her salary exceeds $150,000?

Hint: Use the Chebyshev inequality.

The *sample mean*, designated by $\bar{x}$, is defined by

$$\bar{x} = \sum_{i=1}^{n} x_i / n$$

The *sample variance*, call it s^2, of the data set $x_1, \ldots, x_n$ is defined by

$$s^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 / (n-1)$$

The quantity s, defined by

$$s = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 / (n-1)}$$

is called the *sample standard deviation*.

Chebyshev's Inequality

Let $\bar{x}$ and s be the sample mean and sample standard deviation of the data set consisting of the data $x_1, \ldots, x_n$, where $s > 0$. Let

$$S_k = \{i, 1 \leq i \leq n : |x_i - \bar{x}| < ks\}$$

and let $N(S_k)$ be the number of elements in the set S_k. Then, for any $k \geq 1$,

$$\frac{N(S_k)}{n} \geq 1 - \frac{n-1}{nk^2} > 1 - \frac{1}{k^2}$$

Because Chebyshev's inequality holds universally, it might be expected for given data that the actual percentage of the data values that lie within the interval from $\bar{x} - ks$ to $\bar{x} + ks$ might be quite a bit larger than the bound given by the inequality.

Let $\bar{x}$ and s be the sample mean and sample standard deviation of a data set. Assuming that $s > 0$, Chebyshev's inequality states that for any value of $k \geq 1$, greater than $100(1 - 1/k^2)$ percent of the data lie within the interval from $\bar{x} - ks$ to $\bar{x} + ks$. Thus, by letting $k = 3/2$, we obtain from Chebyshev's inequality that greater than $100(5/9) = 55.56$ percent of the data from any data set lies within a distance $1.5s$ of the sample mean $\bar{x}$; letting $k = 2$ shows that greater than 75 percent of the data lies within $2s$ of the sample mean; and letting $k = 3$ shows that greater than $800/9 \approx 88.9$ percent of the data lies within 3 sample standard deviations of $\bar{x}$.

Suppose now that we are interested in the fraction of data values that exceed the sample mean by at least k sample standard deviations, where k is positive. That is, suppose that $\bar{x}$ and s are the sample mean and the sample standard deviation of the data set $x_1, x_2, \ldots, x_n$. Then, with

$$N(k) = \text{number of } i : x_i - \bar{x} \geq ks$$

what can we say about $N(k)/n$? Clearly,

$$\frac{N(k)}{n} \leq \frac{\text{number of } i : |x_i - \bar{x}| \geq ks}{n}$$

$$\leq \frac{1}{k^2} \quad \text{by Chebyshev's inequality}$$

However, we can make a stronger statement, as is shown in the following one-sided version of Chebyshev's inequality.

The One-Sided Chebyshev Inequality

For $k > 0$,

$$\frac{N(k)}{n} \leq \frac{1}{1 + k^2}$$

23. Of three cards, one is painted red on both sides; one is painted black on both sides; and one is painted red on one side and black on the other. A card is randomly chosen and placed on a table. If the side facing up is red, what is the probability that the other side is also red?

24. A couple has 2 children. What is the probability that both are girls if the eldest is a girl?

25. Fifty-two percent of the students at a certain college are females. Five percent of the students in this college are majoring in computer science. Two percent of the students are women majoring in computer science. If a student is selected at random, find the conditional probability that

    (a) this student is female, given that the student is majoring in computer science;

    (b) this student is majoring in computer science, given that the student is female.

26. A total of 500 married working couples were polled about their annual salaries, with the following information resulting.

| | Husband | |
Wife	Less than $50,000	More than $50,000
Less than $50,000	212	198
More than $50,000	36	54

Thus, for instance, in 36 of the couples the wife earned more and the husband earned less than $50,000. If one of the couples is randomly chosen, what is

    (a) the probability that the husband earns less than $50,000;

    (b) the conditional probability that the wife earns more than $50,000 given that the husband earns more than this amount;

    (c) the conditional probability that the wife earns more than $50,000 given that the husband earns less than this amount?

27. There are two local factories that produce microwaves. Each microwave produced at factory A is defective with probability .05, whereas each one produced at factory B is defective with probability .01. Suppose you purchase two microwaves that were produced at the same factory, which is equally likely to have been either factory A or factory B. If the first microwave that you check is defective, what is the conditional probability that the other one is also defective?

28. A red die, a blue die, and a yellow die (all six-sided) are rolled. We are interested in the probability that the number appearing on the blue die is less than that appearing on the yellow die which is less than that appearing on the red die. (That is, if B (R) [Y] is the number appearing on the blue (red) [yellow] die, then we are interested in $P(B < Y < R)$.)

    (a) What is the probability that no two of the dice land on the same number?

(b) Given that no two of the dice land on the same number, what is the conditional probability that $B < Y < R$?

(c) What is $P(B < Y < R)$?

(d) If we regard the outcome of the experiment as the vector B, R, Y, how many outcomes are there in the sample space?

(e) Without using the answer to (c), determine the number of outcomes that result in $B < Y < R$.

(f) Use the results of parts (d) and (e) to verify your answer to part (c).

29. You ask your neighbor to water a sickly plant while you are on vacation. Without water it will die with probability .8; with water it will die with probability .15. You are 90 percent certain that your neighbor will remember to water the plant.

(a) What is the probability that the plant will be alive when you return?

(b) If it is dead, what is the probability your neighbor forgot to water it?

30. Two balls, each equally likely to be colored either red or blue, are put in an urn. At each stage one of the balls is randomly chosen, its color is noted, and it is then returned to the urn. If the first two balls chosen are colored red, what is the probability that

(a) both balls in the urn are colored red;

(b) the next ball chosen will be red?

31. A total of 600 of the 1,000 people in a retirement community classify themselves as Republicans, while the others classify themselves as Democrats. In a local election in which everyone voted, 60 Republicans voted for the Democratic candidate, and 50 Democrats voted for the Republican candidate. If a randomly chosen community member voted for the Republican, what is the probability that she or he is a Democrat?

32. Each of 2 balls is painted black or gold and then placed in an urn. Suppose that each ball is colored black with probability 1/2, and that these events are independent.

(a) Suppose that you obtain information that the gold paint has been used (and thus at least one of the balls is painted gold). Compute the conditional probability that both balls are painted gold.

(b) Suppose, now, that the urn tips over and 1 ball falls out. It is painted gold. What is the probability that both balls are gold in this case? Explain.

33. Each of 2 cabinets identical in appearance has 2 drawers. Cabinet A contains a silver coin in each drawer, and cabinet B contains a silver coin in one of its drawers and a gold coin in the other. A cabinet is randomly selected, one of its drawers is opened, and a silver coin is found. What is the probability that there is a silver coin in the other drawer?

34. Prostate cancer is the most common type of cancer found in males. As an indicator of whether a male has prostate cancer, doctors often perform a test that measures the level of the PSA protein (prostate specific antigen) that is produced only by the prostate gland. Although higher PSA levels are indicative of cancer, the test is notoriously unreliable. Indeed, the probability that a noncancerous man will have an elevated PSA level is approximately .135, with this probability increasing to approximately .268 if the man does have cancer. If, based on other factors, a physician is 70 percent certain that a male has prostate cancer, what is the conditional probability that he has the cancer given that

      (a) the test indicates an elevated PSA level;

      (b) the test does not indicate an elevated PSA level?

Repeat the preceding, this time assuming that the physician initially believes there is a 30 percent chance the man has prostate cancer.

35. Suppose that an insurance company classifies people into one of three classes — good risks, average risks, and bad risks. Their records indicate that the probabilities that good, average, and bad risk persons will be involved in an accident over a 1-year span are, respectively, .05, .15, and .30. If 20 percent of the population are "good risks," 50 percent are "average risks," and 30 percent are "bad risks," what proportion of people have accidents in a fixed year? If policy holder A had no accidents in 1987, what is the probability that he or she is a good (average) risk?

**36.** A pair of fair dice is rolled. Let E denote the event that the sum of the dice is equal to 7.

      (a) Show that E is independent of the event that the first die lands on 4.

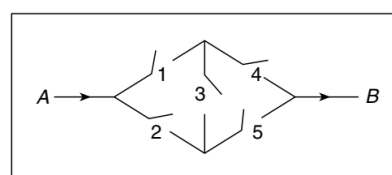   **(b)** Show that E is independent of the event that the second die lands on 3.

37. The probability of the closing of the ith relay in the circuits shown is given by p_i, i = 1, 2, 3, 4, 5. If all relays function independently, what is the probability that a current flows between A and B for the respective circuits?
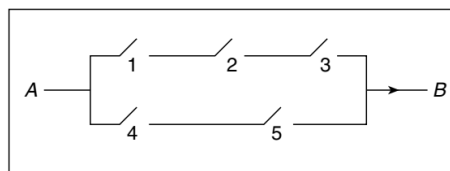
(a)
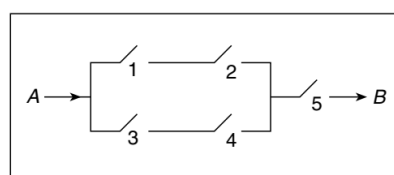
(b)   (c)

38. An engineering system consisting of n components is said to be a k-out-of-n system ($k \leq n$) if the system functions if and only if at least k of the n components function. Suppose that all components function independently of each other.

> **(a)** If the ith component functions with probability P_i, $i = 1, 2, 3, 4$, compute the probability that a 2-out-of-4 system functions.

> **(b)** Repeat (a) for a 3-out-of-5 system.

39. Five independent flips of a fair coin are made. Find the probability that

> **(a)** the first three flips are the same;

> **(b)** either the first three flips are the same, or the last three flips are the same;

> **(c)** there are at least two heads among the first three flips, and at least two tails among the last three flips.

40. Suppose that n independent trials, each of which results in any of the outcomes 0, 1, or 2, with respective probabilities .3, .5, and .2, are performed. Find the probability that both outcome 1 and outcome 2 occur at least once. (*Hint:* Consider the complementary probability.)

41. A parallel system functions whenever at least one of its components works. Consider a parallel system of n components, and suppose that each component independently works with probability 1/2. Find the conditional probability that component 1 works, given that the system is functioning.

42. A certain organism possesses a pair of each of 5 different genes (which we will designate by the first 5 letters of the English alphabet). Each gene appears in 2 forms (which we designate by lowercase and capital letters). The capital letter will be assumed to be the dominant gene in the sense that if an organism possesses the gene pair xX, then it will outwardly have the appearance of the X gene. For instance, if X stands for brown eyes and x for blue eyes, then an individual having either gene pair XX or xX will have brown eyes, whereas one having gene pair xx will be blue-eyed. The characteristic appearance of an organism is called its *phenotype*, whereas its genetic constitution is called its *genotype*. (Thus 2 organisms with respective genotypes *aA, bB, cc, dD, ee* and *AA, BB, cc, DD, ee* would have different genotypes but the same phenotype.) In a mating between 2 organisms each one contributes, at random, one of its gene pairs of each type. The 5 contributions of an organism (one of each of the 5 types) are assumed to be independent and are also independent of the contributions of its mate. In a mating between organisms having genotypes *aA, bB, cC, dD, eE*, and *aa, bB, cc, Dd, ee*, what is the probability that the progeny will (1) phenotypically, (2) genotypically resemble

> **(a)** the first parent;

> **(b)** the second parent;

> **(c)** either parent;

> **(d)** neither parent?

43. Three prisoners are informed by their jailer that one of them has been chosen at random to be executed, and the other two are to be freed. Prisoner *A* asks the jailer to tell him privately which of his fellow prisoners will be set free, claiming that there would be no harm in divulging this information because he already knows that at least one of the two will go free. The jailer refuses to answer this question, pointing out that if *A* knew which of his fellow prisoners were to be set free, then his own probability of being executed would rise from 1/3 to 1/2 because he would then be one of two prisoners. What do you think of the jailer's reasoning?

44. Although both my parents have brown eyes, I have blue eyes. What is the probability that my sister has blue eyes? (As stated in Problem 42, an individual who receives a blue-eyed gene from each parent will have blue eyes, whereas one who receives one blue-eyed and one brown-eyed gene will have brown eyes.)

45. In a 7 game series played with two teams, the first team to win a total of 4 games is the winner. Suppose that each game played is independently won by team *A* with probability *p*.

(a) Given that one team leads 3 to 0, what is the probability that it is team *A* that is leading?

(b) Given that one team leads 3 to 0, what is the probability that team wins the series?

(c) If *p* = 1/2, what is the probability that the team that wins the first game wins the series?

46. Suppose that distinct integer values are written on each of 3 cards. Suppose you are to be offered these cards in a random order. When you are offered a card you must immediately either accept it or reject it. If you accept a card, the process ends. If you reject a card, then the next card (if a card remains) is offered. If you reject the first two cards offered, then you must accept the final card.

(a) If you plan to accept the first card offered, what is the probability that you will accept the highest valued card?

(b) If you plan to reject the first card offered, and to then accept the second card if and only if its value is greater than the value of the first card, what is the probability that you will accept the highest valued card?

47. Let *A*, *B*, *C* be events such that *P(A)* = .2, *P(B)* = .3, *P(C)* = .4. Find the probability that at least one of the events *A* and *B* occurs if

(a) *A* and *B* are mutually exclusive;

(b) *A* and *B* are independent.

Find the probability that all of the events *A*, *B*, *C* occur if

(c) *A*, *B*, *C* are independent;

(d) *A*, *B*, *C* are mutually exclusive.

48. Two percent of woman of age 45 who participate in routine screening have breast cancer. Ninety percent of those with breast cancer have positive mammographies. Ten percent of the women who do not have breast cancer will also have positive mammographies. Given a woman has a positive mammography, what is the probability she has breast cancer?

49. Twelve percent of all US households are in California. A total of 3.3 percent of all US households earn over 250, 000 per year, while a total of 6.3 percent of California households earn over 250, 000 per year. If a randomly chosen US household earns over 250, 000 per year, what is the probability it is from California?

50. There is a 60 percent chance that the event A will occur. If A does not occur, there is a 10 percent chance that B will occur. What is the probability that at least one of the events A or B occur?

51. Suppose distinct values are written on each of three cards, which are then randomly given the designations A, B, and C. The values on cards A and B are then compared. What is the probability that the smaller of these values is also smaller than the value on card C?